Bounds for the Singular Values of a Matrix with Nonnegative Eigenvalues

Izchak Lewkowicz
Department of Electrical Engineering
Technion — Israel Institute of Technology
Haifa 32000, Israel

Submitted by Emeric Deutsch

ABSTRACT

An upper bound and a lower bound for *each* singular value of a matrix with *nonnegative* eigenvalues are derived. These bounds are based upon the matrix spectral decomposition. It is shown that this estimate for each singular value is tighter than a well known one, based upon the condition number of the eigenvector matrix. Note, however, that the known estimate is also applicable to matrices with *complex* eigenvalues. A property of projection matrices, used in the proof, is discussed as well.

I. INTRODUCTION

Several estimates for *each* singular value of a given matrix are available, e.g. [1,2]. The best known is the following from [1]: Let $A \in \mathbb{C}^{n\times n}$ be a nondefective matrix, i.e. with n linearly independent eigenvectors, forming a matrix $V \in \mathbb{C}^{n\times n}$ such that $A = V\{\mathrm{diag}[\lambda_1(A),\dots,\lambda_n(A)]\}V^{-1}$, where the eigenvalues of A, $\lambda_k(A) \in \mathbb{C}$, $1 \leqslant k \leqslant n$, are ordered, hereafter, so that $|\lambda_1(A)| \geqslant \cdots \geqslant |\lambda_n(A)|$. If we denote by $\sigma_1(A) \geqslant \cdots \geqslant \sigma_n(A)$ the singular values of A, and by $\chi(V) \triangleq \|V\|_2\|V^{-1}\|_2$ the *condition number* of the nonsingular eigenvector matrix V, then the following estimate for each singular value holds:

$$\frac{|\lambda_k(A)|}{\chi(V)} \leqslant \sigma_k(A) \leqslant |\lambda_k(A)|\chi(V) \qquad \forall k. \tag{1}$$

In this note, a new estimate for each singular value of a matrix with nonnegative eigenvalues is introduced, and is compared with (1).

655 Avenue of the Americas, New York, NY 10010 0024-3795/89/$3.50

II. PROPERTIES OF PROJECTION MATRICES

We establish now some notation concerning projection matrices. First, for a nondefective matrix, we have the following spectral decomposition (e.g. [3]):

$$A = \sum_{k=1}^{n} \lambda_k(A) V_{(:,k)} V^{-1}_{(k,:)} \triangleq \sum_{k=1}^{n} \lambda_k(A) P_k, \tag{2}$$

where $V_{(:,k)}$ is the kth column vector in the matrix V, and $V^{-1}_{(k,:)}$ is the kth row vector in the matrix V^{-1}. Each P_k is a *not necessarily orthogonal* projection matrix of rank 1; i.e., for $1 \leqslant i, j \leqslant n$, (i) $P_i P_j = P_i \delta_{ij}$, where δ_{ij} is the Kronecker δ (ii) $\|P_i\|_2 \geqslant 1$, and (iii) $\sum_{i=1}^{n} P_i = I$.

Next, we denote by N the set of indices $N = \{1, 2, \ldots, n\}$, and by J, any *nonempty proper* subset of N. Using this notation, we have that for all J,

$$P(J) \triangleq \sum_{k \in J} P_k$$

is a projection matrix too, and so is

$$I - P(J) \triangleq \sum_{k \in N - J} P_k.$$

The following lemma describes the relation among the singular values of a projection matrix $P(J)$ and those of $I - P(J)$.

LEMMA 1. *Given a projection matrix $P(J) \in \mathbb{C}^{n \times n}$, where J has m elements $1 \leqslant m < n$. Then, the following relations hold:*

$$\begin{aligned} &\sigma_k[P(J)] = \sigma_k[I - P(J)] \geqslant 1, && 1 \leqslant k \leqslant m, \\ &\sigma_k[P(J)] = 0, \quad \sigma_k[I - P(J)] = 1, && m + 1 \leqslant k \leqslant n - m, \\ &\sigma_k[P(J)] = \sigma_k[I - P(J)] = 0, && n - m + 1 \leqslant k \leqslant n. \end{aligned} \tag{3}$$

Proof. We start by constructing a unitary similarity transformation identical to the one used in Schur's triangularization process. The $n - m$ vectors $V_{(:,k)}$, $k \in N - J$, span the *kernel* subspace of $P(J)$. Using the

Gram-Schmidt process, we form from it an orthonormal basis for the same subspace, and denote its elements by $U_{(:,i)}$, $m+1 \leqslant i \leqslant n$.

Now, the m vectors $V_{(:,k)}$, $k \in J$ span the *image* subspace of $P(J)$. Using again the Gram-Schmidt process, we complete the orthonormal basis $U_{(:,i)}$, $1 \leqslant i \leqslant m$, for the entire n dimensional vector space. Arranging these n orthonormal vectors in a unitary matrix U results in

$$U^*P(J)U = \begin{bmatrix} I_m & 0 \\ \tilde{P}(J) & 0 \end{bmatrix}, \tag{4}$$

where I_m is the unit matrix of order m, $(\cdot)^*$ is the complex conjugate transpose of $(\cdot)$, and $\tilde{P}(J) \in \mathbb{C}^{(n-m)\times m}$. Also,

$$U^*[I - P(J)]U = \begin{bmatrix} 0 & 0 \\ -\tilde{P}(J) & I_{n-m} \end{bmatrix}. \tag{5}$$

Next, from (4) we have that

$$U^*P^*(J)P(J)U = \begin{bmatrix} I_m + \tilde{P}^*(J)\tilde{P}(J) & 0 \\ 0 & 0 \end{bmatrix},$$

and similarly, from (5)

$$U^*[I - P(J)][I - P^*(J)]U = \begin{bmatrix} 0 & 0 \\ 0 & I_{n-m} + \tilde{P}(J)\tilde{P}^*(J) \end{bmatrix}.$$

Now, using the fact that [4, p. 318] for any matrix $M \in \mathbb{C}^{n\times m}$, $n \geqslant m$, for each k, $1 \leqslant k \leqslant m$,

$$\sigma_k^2(M) = \lambda_k(M^*M) = \lambda_k(MM^*),$$

and for $m+1 \leqslant k \leqslant n$, $\sigma_k(M) = 0$, the proof of Lemma 1 can be completed. ∎

We recall that [4, p. 321] for every matrix $M \in \mathbb{C}^{n\times n}$, $\sigma_1(M) = \|M\|_2$. Then, Lemma 1 implies in particular that $\forall J$

$$\|P(J)\|_2 = \|I - P(J)\|_2. \tag{6}$$

In fact, the last equality holds for every projection operator P in a Hilbert space, whenever both P and $I-P$ are not projectors to the null space only. This observation is closely related to the work in [5].

III. AN ESTIMATE FOR EACH SINGULAR VALUE

We denote by $\alpha(V)$ the maximal norm of the projection matrices, i.e.

$$\alpha(V) \triangleq \max_{J \subset N} \|P(J)\|_2,$$

where due to (6), the search for $\alpha(V)$ can be carried out over all J's containing $n/2$ elements at the most. At this stage, we can establish the following upper bound for the 2-norm of a matrix with nonnegative eigenvalues.

LEMMA 2. *Given a nondefective matrix* $A \in \mathbb{C}^{n\times n}$, *with the spectral decomposition* (2). *If in addition* $\forall k$, $1 \leqslant k \leqslant n$, $\lambda_k(A) \geqslant 0$, *then the following relation holds*:

$$\|A\|_2 \leqslant \lambda_1(A)\alpha(V). \tag{7}$$

Proof. In view of (6), we can select a set of indices J_0 so that $\alpha(V) = \|P(J_0)\|_2$ and

$$\left\| \sum_{k\in J_0} \lambda_k(A)P_k \right\|_2 \geqslant \left\| \sum_{k \in N - J_0} \lambda_k(A)P_k \right\|_2 .$$

Note that if X and Y are matrices with compatible sizes, then

$$\|X+Y\| \geqslant \|X\| \geqslant \|Y\| \quad \Rightarrow \quad \|X+Y\| \geqslant \|X+\theta Y\| \geqslant \|X\|, \quad 1 \geqslant \theta \geqslant 0, \tag{8}$$

$$\|X\| \geqslant \|Y\| \geqslant \|X+Y\| \quad \Rightarrow \quad \|X\| \geqslant \|X+\theta Y\| \geqslant \|X+Y\|, \quad 1 \geqslant \theta \geqslant 0. \tag{9}$$

From (8) and the definition of $\alpha(V)$, we have that $\forall \theta_k$, $0 \leqslant \theta_k \leqslant 1$, and $k \in J_0$,

$$\left\| \sum_{k \in J_0} \theta_k P_k \right\|_2 \leqslant \alpha(V). \tag{10}$$

Multiplying both sides of (10) by $\lambda_1(A)$ yields that for each $k, k \in J_0$, we can select a θ_k, $0 \leqslant \theta_k \leqslant 1$, so that

$$\left\| \sum_{k \in J_0} \lambda_k(A) P_k \right\|_2 = \left\| \sum_{k \in J_0} [\lambda_1(A)\theta_k] P_k \right\|_2 \leqslant \lambda_1(A)\alpha(V). \tag{11}$$

Using (11) and again the definition of $\alpha(V)$, together with (9), yields that $\forall \theta_k$, $k \in N - J_0$ and $0 \leqslant \theta_k \leqslant 1$, the following relation holds:

$$\|A\|_2 \leqslant \left\| \sum_{k \in J_0} \lambda_k(A) P_k + \sum_{k \in N - J_0} \lambda_k(A)\theta_k P_k \right\|_2 \leqslant \left\| \sum_{k \in J_0} \lambda_k(A) P_k \right\|_2. \tag{12}$$

A sufficient condition for the left inequality to hold with equality is $\theta_k = 1$ $\forall k$, $k \in N - J_0$, and a sufficient condition for the right inequality to hold with equality is $\theta_k = 0$ $\forall k$, $k \in N - J_0$. From (11) and (12), the proof is completed. ■

The following main result provides us with an estimate for *each* singular value of a given nondefective matrix with nonnegative eigenvalues.

THEOREM. *Given a nondefective matrix* $A \in \mathbb{C}^{n \times n}$ *with the spectral decomposition* (2). *If in addition* $\forall k$, $1 \leqslant k \leqslant n$, $\lambda_k(A) \geqslant 0$, *then the following relations hold*:

$$\frac{\lambda_k(A)}{\alpha(V)} \leqslant \sigma_k(A) \leqslant \lambda_k(A)\alpha(V). \tag{13}$$

Proof. First, we prove the right inequality. Due to Lemma 2, the claim is established for $k = 1$. Next, we recall that (e.g. [1]) for every two matrices M, N with compatible sizes, $\forall k$, $1 \leqslant k \leqslant n$, $\sigma_k(M) - \sigma_1(N) \leqslant \sigma_k(M - N)$. Taking $M = A$ and $N = \sum_{j=k}^{n} \lambda_j(A) P_j$ yields that $\forall k$, $1 \leqslant k \leqslant n$,

$$\sigma_k(A) - \left\| \sum_{j=k}^{n} \lambda_j(A) P_j \right\|_2 \leqslant \sigma_k\left[A - \sum_{j=k}^{n} \lambda_j(A) P_j \right]. \tag{14}$$

Now, for every matrix $M \in \mathbb{C}^{n \times n}$, $\forall k$, $2 \leqslant k \leqslant n$, $\operatorname{rank}(M) \leqslant k - 1$ is equivalent to $\sigma_k(M)=0$. Hence the right-hand side of (14) vanishes. Hence, applying (7), we conclude that, $\forall k$, $2 \leqslant k \leqslant n$,

$$\sigma_k(A) \leqslant \left\| \sum_{j=k}^{n} \lambda_j(A) P_j \right\|_2 \leqslant \lambda_k(A) \left\| \sum_{j=k}^{n} P_j \right\|_2 \leqslant \lambda_k(A)\alpha(V).$$

This establishes the right inequality in (13).

For the proof of the left inequality in (13), we first assume that A is nonsingular. Now, we consider the following facts $\forall k$, $1 \leqslant k \leqslant n$: (i) from (2), $\lambda_k(A^{-1}) = \lambda_{n+1-k}^{-1}(A)$ and $P_k(A^{-1}) = P_{n+1-k}(A)$, and (ii), $\sigma_k(A^{-1}) = \sigma_{n+1-k}^{-1}(A)$. Combining the right inequality in (13) with the last two facts establishes the left inequality for a nonsingular matrix.

Now, if A is singular (i.e., for some m, $2 \leqslant m \leqslant n$, we have $\lambda_k(A)=0$ for $k \geqslant m$), then, using (2), we can always define, for some $\varepsilon > 0$, the following nonsingular matrix $\tilde{A}$:

$$\tilde{A} \triangleq \sum_{k=1}^{m-1} \lambda_k(A) P_k + \sum_{k=m}^{n} \varepsilon P_k.$$

First, the left inequality of (13) is obtained for $\tilde{A}$. Then, since each singular value is a continuous function of the matrix elements, taking the limit $\varepsilon \to 0$, the proof is completed. ■

Note that the theorem holds for a nondefective matrix with *nonpositive* eigenvalues as well, but not for a matrix with *both* positive and negative real eigenvalues.

Next, it is shown that the estimates in the theorem are *tighter* than those from [1], quoted in (1). We introduce here the following notation: $I(J) \triangleq \operatorname{diag}\{\delta_{1j}, \delta_{2j}, \ldots, \delta_{nj}\}$, $j \in J$. From this notation we have that for each J,

$$P(J) = VI(J)V^{-1}.$$

Taking the 2-norm over the last relation for $J = J_0$ yields

$$\alpha(V) \leqslant \chi(V), \tag{15}$$

where we have used the fact that $\forall J$, $\|I(J)\|_2 = 1$. Consequently, we have that the estimates in the theorem are tighter than those in (1), which in turn hold with equality if the matrix A is normal.

IV. DISCUSSION

In this section several aspects of the new estimate are examined. First, we present a generalization of the theorem.

Given a matrix $\hat{A} \in \mathbb{C}^{n\times n}$, with any eigenvalues $\lambda_k(\hat{A}) \in \mathbb{C}$, $1 \leqslant k \leqslant n$. Assume, without loss of generality, that $\hat{A}$ is given in its Schur triangular form. Next, define the following unitary, diagonal matrix $U \in \mathbb{C}^{n\times n}$:

$$u_{ij} \triangleq \begin{cases} 0, & i \neq j, \\ \hat{a}_{ii}^*/|\hat{a}_{ii}|, & \hat{a}_{ii} \neq 0, \\ 1, & \hat{a}_{ii} = 0, \end{cases}$$

and denote by A the following triangular matrix associated with $\hat{A}$:

$$A \triangleq \hat{A}U.$$

In view of the fact that $\forall k$ $\lambda_k(A) = \|\lambda_k(\hat{A})\|$, we can state the following corollary of the main result.

COROLLARY. *Given a matrix $\hat{A} \in \mathbb{C}^{n\times n}$, with eigenvalues $\lambda_k(\hat{A}) \in \mathbb{C}$, $1 \leqslant k \leqslant n$. If the matrix A associated with it is nondefective with the spectral decomposition* (2), *then $\forall k$, $1 \leqslant k \leqslant n$, the following relations hold*:

$$\frac{\lambda_k(A)}{\alpha(V)} \leqslant \sigma_k(\hat{A}) \leqslant \lambda_k(A)\alpha(V). \tag{16}$$

Proof. Due to the definitions, we have that $\forall k$, $1 \leqslant k \leqslant n$, the following properties hold: (i) $\lambda_k(A) \geqslant 0$ and (ii) $\sigma_k(\hat{A}) = \sigma_k(A)$. Using the theorem, the proof is established. ■

Although a generalization of the theorem, the result in the corollary suffers from two drawbacks. First, unlike the result from [1], quoted in (1), the corollary is not applicable to *all* the family of nondefective matrices, e.g.

$$\hat{A} = \begin{bmatrix} -1 & 0 \\ 1 & 1 \end{bmatrix},$$

since the associated matrix A turns out to be defective.

The second drawback of the corollary is that although the estimates in the theorem are tighter than those from [1], quoted in (1), this desired property is not necessarily valid for the estimates in the corollary. Even when the matrix A associated with a given $\hat{A}$ is nondefective, $\alpha(V)$ is not necessarily smaller than $\chi(\hat{V})$, the condition number of the eigenvector matrix of $\hat{A}$. Consider the following example:

$$\hat{A} = \begin{bmatrix} 1 & 0 \\ 1 & -2 \end{bmatrix}.$$

For simplicity we consider just $\sigma_1(\hat{A}) = 2.29$. Then, (1) provides us with $|\lambda_1(\hat{A})|\chi(\hat{V}) = 2.79$ as an upper bound, while the corollary yields $\lambda_1(A)\alpha(V) = 2.83$ only, where for conformity, each eigenvector is normalized so that $\forall k$ $\|V_{(:,k)}\|_\infty = 1$.

This note is concluded with the following example, which illustrates the fact that it may be difficult to formulate another estimate for each singular value, based upon the spectral decomposition, which is even tighter than the one in the theorem. For simplicity, we consider only the special case of Lemma 2.

Since from the definition of $\alpha(V)$

$$\max_{1 \leqslant k \leqslant n} \|P_k\|_2 \leqslant \alpha(V), \tag{17}$$

one might think that in the same conditions, Lemma 2 can be replaced by the following:

$$\|A\|_2 \leqslant \lambda_1(A) \max_{1 \leqslant k \leqslant n} \|P_k\|_2. \tag{18}$$

Indeed, for $n \leqslant 3$, (17) holds with equality and hence (18) is valid. However, for $n \geqslant 4$, (18) does not necessarily hold; e.g.

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 2 & 2 & 5 & 0 \\ -2 & -2 & 0 & 5 \end{bmatrix}, \qquad \|A\|_2 = 6.34.$$

Lemma 2 yields $\lambda_1(A)\alpha(V) = 7.07$ as an upper bound for $\|A\|_2$, but (18) results in $\lambda_1(A)\max_{1 \leqslant k \leqslant n}\|P_k\|_2 = 6.12$, which fails as a bound.

We also note that from (1), $\lambda_1(A)\chi(V) = 10.84$, which is more conservative than the new bound, due to Lemma 2.

I wish to thank Professor R. Sivan from the Department of Electrical Engineering and Professor A. Berman from the Department of Mathematics, *both at the Technion, who helped me in my Master's Thesis, which also led to this note.*

REFERENCES

1 A. Ruhe, On the closeness of eigenvalues and SI values for almost normal matrices, *Linear Algebra Appl.* 11:87–94.
2 L. Qi, Some simple estimates for the singular values of a matrix, *Linear Algebra Appl.* 56:105–119 (1984).
3 K. Hoffman and R. Kunze, *Linear Algebra*, 2nd ed., Prentice-Hall, Englewood Cliffs, N.J., 1971.
4 G. W. Stewart, *Introduction to Matrix Computations*, Academic, New York, 1973.
5 S. N. Afriat, Orthogonal and oblique projectors and the characteristics of pairs of vector spaces, *Proc. Cambridge Philos. Soc.*, 1957, pp. 800–816.

Received 16 February 1987; final manuscript accepted 27 April 1988